Pergamon

PII: S0042-6989(97)00215-0

Vision Res., Vol. 38, No. 6, pp. 853–863, 1998

Printed in Great Britain
0042-6989/98 $19.00 + 0.00

A Global Process in Motion Segregation

MARY J. BRAVO*

Received 24 July 1996; in revised form 23 May 1997

Observers viewed sparse random dot cinematograms in which the moving dots were confined to eight windows. The motions in seven of the windows were consistent with a global flow pattern, while the direction of motion in the eighth window deviated from this pattern. The observer's task was to determine which of the eight windows contained the inconsistent motion. The task was performed on two types of global flow patterns: spirals, which appear rigid, and deformations, which appear highly non-rigid. Although these patterns produce qualitatively different global percepts, they are exactly matched in their local velocities and velocity differences. Observers were better able to locate an inconsistent motion in spiral patterns than in deformation patterns, indicating that they were using more than just local motion information to find the target. This result is taken as indirect support for a segregation process that involves fitting the stimulus with a global motion pattern and segregating motions inconsistent with this pattern. © 1998 Published by Elsevier Science Ltd. All rights reserved.

Optic flow Image segmentation

INTRODUCTION

One simple but powerful assumption that can be made about the retinal motions produced by surfaces is that velocity varies smoothly across a surface and so a discontinuity in velocity indicates a surface boundary (Marr, 1982). The human visual system evidently incorporates this assumption into motion processing: under appropriate conditions a velocity discontinuity appears as a perceptually salient edge. Moreover, these motion-defined edges can serve the same functions as contrast edges in 2D shape perception (Narwot, Shannon & Rizzo, 1996; Banton & Levi, 1993), spatial illusions (Cavanagh, 1989) apparent motion, (Petersik, Hicks & Pantle, 1978), and stereopsis (Halpern, 1991).

While it is clear that velocity discontinuities provide reliable information about surface boundaries, it is also clear that such discontinuities are, by themselves, insufficient to organize a scene. Figure 1(a) depicts the motions generated by a surface moving in front of a second surface. A segregation process based solely on local velocity differences would divide this stimulus into three regions. A human observer, however, would associate the outer regions and perceive two surfaces. Figure 1(b) depicts the patchy flow field that an observer might see when approaching a partially occluded wall. Although the neighboring patches in this flow field have very different velocities, they all originate from the same surface. A segregation process that detects velocity differences would parse this scene into eight small regions, while a human observer would probably see a single extended surface.

Such considerations have led to the proposal of a second, global approach for using motion information to organize a scene. This second approach involves fitting the stimulus with a limited set of motion patterns that correspond to the flow fields an active observer commonly encounters [e.g., the expanding pattern depicted in Fig. 1(b)]. If many of the local motions in a scene can be well fit by a single motion pattern, then these motions are grouped together and segregated from motions that are inconsistent with the pattern. This idea, which was initially developed for computer vision (see, for example, Bergen, Burt, Hingorani & Peleg, 1992; Irani, Rousso & Peleg, 1992; Darrell & Pentland, 1991; Black & Anandan, 1993), has recently been incorporated into theories of human vision (Yuille & Grzywacz, 1997).

There is abundant evidence that humans are quite sensitive to certain flow patterns (Lappin, Norman & Mowafy, 1991; Freeman & Harris, 1992; Morrone, Burr & Vaina, 1995) and that we can use these patterns to derive heading information (Warren & Hannon, 1988; Warren, Morris & Kalish, 1988). In addition, there are cells in the medial superior temporal area (MST) of the monkey brain which respond selectively to these same global flow patterns (Saito, Yukie, Tanaka, Hikosaka, Fukuda & Iwai, 1986; Duffy & Wurtz, 1991b; Lagae, Maes, Raiguel, Xiao & Orban, 1994; Graziano, Andersen & Snowden, 1994). It is not known, however, whether humans can use these global flow patterns to segregate motions produced by different surfaces. The goal of this study was to look for such evidence.

To determine whether humans can use global motion

*Department of Psychology, University of Pennsylvania, 3815 Walnut St, Philadelphia, PA 19104-6228, U.S.A.

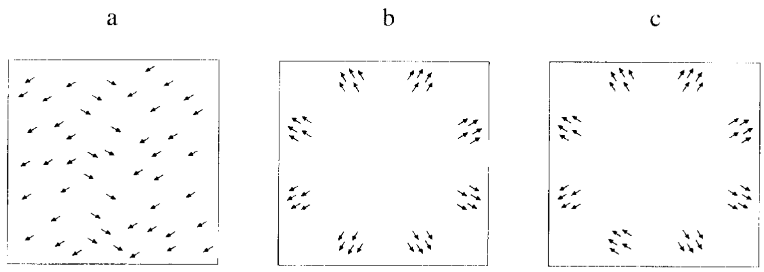

FIGURE 1. (a) Flow field produced by two surfaces translating at different velocities relative to the observer. (b) Flow field seen as an observer approaches a partially occluded wall. (c) Cartoon of the stimuli used here (target in lower left).

patterns to segregate an image, I measured how well observers could locate a target that is moving inconsistently with a global flow pattern. Such a stimulus will ordinarily produce velocity discontinuities in the vicinity of the target. Given that these velocity discontinuities provide powerful information for segregation, the key to this study is to both reduce and control for the information provided by these local differences in velocity.

To reduce the efficacy of velocity discontinuities for segregation, motion was confined to windows as shown in Fig. 1(c). If the visual system computes velocity differences locally, then this windowing will obscure the differences surrounding the target. If gradients are computed across neighboring windows, then, as noted above, the visual system will likely detect sizable differences throughout the display and not just in the vicinity of the target.

While windowing reduces the efficacy of velocity differences for segregation, it may not completely eliminate this source of information. When neighboring windows are compared, the largest velocity differences in the display may still surround the target. To control for any residual effects of velocity differences, it is necessary to compare performance on two global flow patterns that are composed of identical local velocities and local velocity differences. As described below, two stimuli that are so matched are spirals (combinations of expansions and rotations) and deformations (expansions along one axis and contractions along the orthogonal axis). Despite their dissimilar global appearance, spiral and deformation patterns have identical local motion characteristics. Thus, if observers base segregation solely on local motion differences, they should be able to find the target motion equally well in either pattern.

MAIN EXPERIMENT

Methods

Apparatus. Random dot movies were displayed at 100 Hz on a Tektronix 608 monitor (P31 phosphor) controlled by a Strawberry Tree D/A board and a Macintosh computer. At the viewing distance of 57.3 cm, the display's spatial resolution was 15 sec of arc. Subjects viewed the monitor monocularly through a reduction tube which included a 1 log unit neutral density filter to reduce the visible persistence of the phosphor. The only light visible to the subject came from the dots in the eight windows of the display and from the fixation mark. The luminance of the moving dot texture (47 dots/cm^2) was 3 cd/m^2.

Display. The dots were confined to eight square windows (0.8 deg.v.a. wide) arranged in a circle. The circle was centered on the fixation mark and had a radius of 4 deg. v.a. Taking 0 deg to be the 3 o'clock position on the circle, the first window was centered at 22.5 deg and the remaining windows were evenly spaced at 45 deg intervals around the circle [Fig. 2(a)]. The eight windows contained 30 randomly positioned dots apiece. On each frame of the movie, the dots within a window were all displaced by a constant amount in a constant direction. When a dot reached the edge of a window it wrapped around to the opposite edge. All the dots in the stimulus moved at 2.5 deg/sec* only the direction of dot motion varied across windows. Three methods were used to

*Because of the finite spatial resolution of the display system, speed varied slightly with the direction of motion, however, this variation was well under 5% when measured over a 100 msec interval.

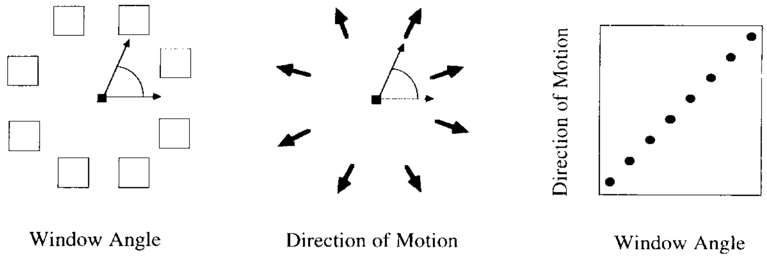

FIGURE 2. (a) Window configuration and definition of window angle. (b) Definition of direction of motion. (c) Plot of direction of motion against window angle for an expansion pattern.

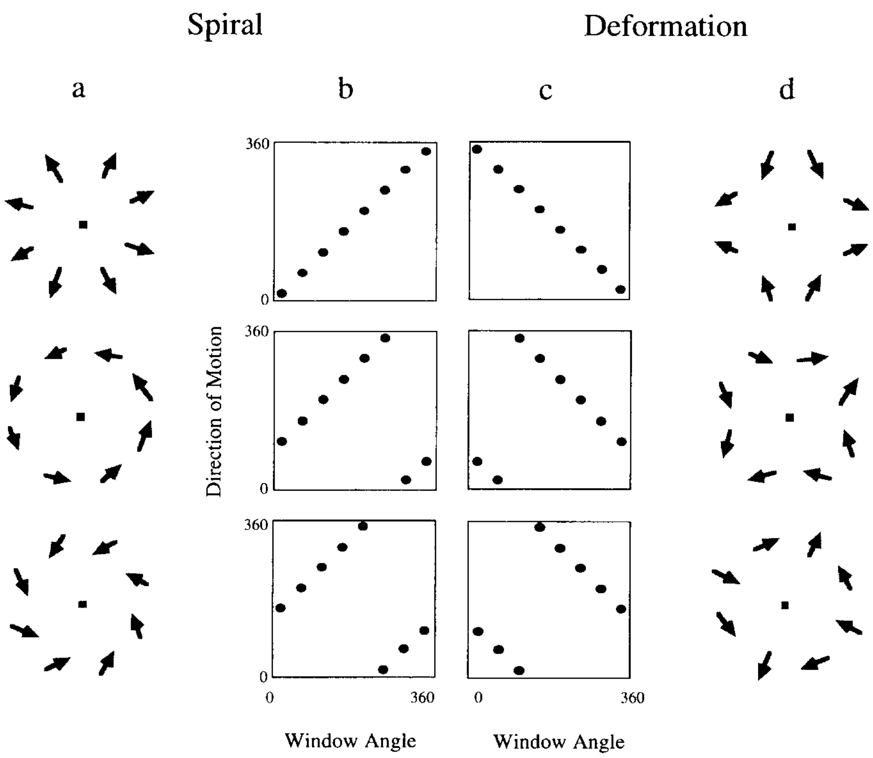

FIGURE 3. (a) Pattern of directions for spiral stimuli. The expansion pattern (top) and rotation pattern (middle) are special cases of the spiral stimuli; the majority of the stimuli were combinations of these two patterns (bottom). (b) For all spiral stimuli, as one moves around the circle of windows, direction of motion changes by the same angular amount. Thus, when direction of motion is plotted against window angle, it produces a line of slope 1. Spiral stimuli differed only in the location of the y-intercept of this line. (c) For all deformation stimuli, direction of motion and window angle changed by the same angular amount but in the opposite direction. For these stimuli, a plot of direction of motion against window angle produces a line of slope -1. (d) Corresponding pattern of directions for deformation stimuli.

assign directions to the windows, with each method producing a different global flow pattern.

Spiral

On each trial, the first window was assigned a random direction between 0 and 359 deg. This direction was then increased by 45 deg and assigned to the second window. The direction was increased again by 45 deg and assigned to the third window, and so on. Recall that the windows were spaced at 45-deg intervals around the circle, so moving from one window to the next results in a change in the direction of motion that was equal to the change in the window angle. This relationship between direction of motion and window angle is shown graphically in Fig. 2. As Fig. 3(a and b) indicate, this method of assigning directions to the windows produces a range of spiral patterns centered on the fixation mark. The relative amounts of expansion and rotation in the spiral pattern are determined by the starting direction: when the direction of motion and the window angle are equal, a pure expansion results, when the direction of motion and the window angle differ by 90 deg, a pure rotation is produced.

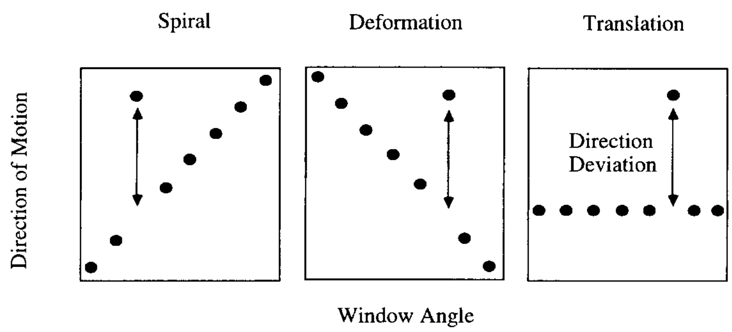

FIGURE 4. A target motion was defined by randomly selecting one window and changing its direction of motion by a variable amount. This amount is referred to as the direction deviation of the target. Note that the changes in direction across windows are identical in magnitude for the spiral and deformation stimuli; it is only the sign of these changes that differs across the two conditions.

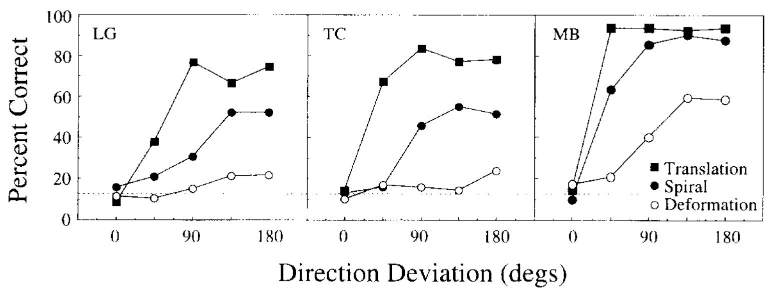

FIGURE 5. Percentage of trials in which the subject correctly identified the target window plotted against the direction deviation of the target. Observers had a 1 in 8 chance of guessing the target window, and this lower limit on performance is indicated by the dashed line. Each panel shows the performance of one subject on the three conditions: with filled squares representing the translation condition, filled circles the spiral condition, and open circles the deformation condition. Note the large discrepancy in performance on the spiral and deformation conditions.

Deformation

On every trial, a randomly selected direction was assigned to the first window, just as in the spiral stimuli. This direction was then *decreased* by 45 deg and assigned to the second window. The direction was decreased again by 45 deg and assigned to the third window, and so on. Thus, in deformation stimuli, as one moves 45 deg around the circle of windows, the direction of motion changes by the same angle but with the opposite sign. Figure 3(c) shows this relationship between window angle and direction of motion, and Fig. 3(d) shows the resulting deformation patterns. Note that the change in direction across neighboring windows is identical for the spiral and deformation stimuli, it is only the sign of this change that differs between the two stimuli.

Translation

On every trial, a direction was selected randomly from 0 to 359 deg, and this direction was assigned to all windows to produce a global translation. Thus, unlike spiral and deformation conditions in which the motion in neighboring windows differed by 45 deg, in the translation condition neighboring windows have the same direction of motion. Consequently, the translation condition differs from the spiral and deformation conditions in both its global pattern and its local motion characteristics. Although this condition cannot be used to differentiate between local and global segregation processes, it likely indicates the upper limit to performance on this task.

Into each of these three types of global flow patterns, a target motion was inserted. One window was selected at random and the direction of motion within that window was changed by either 0, ±45, ±90, ±135, or 180 deg (Fig. 4). The speed of the target motion was unchanged.

There are two characteristics of these stimuli that are worth noting. The first is that an observer could not examine a single window and determine whether it was the target. There was no correlation between target window and target direction: the target could appear in any window and have any direction of motion between 0 and 359 deg. Similarly, the direction of motion in the non-target windows varied between 0 and 359 deg. Thus, the observer could only locate the target by comparing the motions in several windows.

Secondly, the motion within a window was always a translation, and so there was a discrepancy between the local flow and the global flow in the spiral and deformation stimuli. In a true rotating flow pattern, for example, the dots within the windows would move along curved paths and neighboring dots would have slightly different velocities. In the rotating flow pattern used here, all of the dots within a window moved with a constant velocity. This constant velocity was equal to the velocity appropriate for a dot located at the center of the window. Thus, dots at the edges of the windows deviated the most from true flow, but the magnitude of this deviation was identical for all spiral and deformation stimuli. Of course, whether or not this approximation to true flow will affect performance depends critically on window size and eccentricity. A pilot study indicated that under the conditions of this experiment, substituting translations for local flow had no effect on performance and was imperceptible to the observers. The advantage of using local translations is that it guarantees that the local motion properties of these stimuli would be the same across the different global patterns.*

Procedure. During the first 20 min of an initial training

*A second consequence of substituting local translations for true local flow is that it prevents observers from organizing these stimuli using the similarity of higher-level motion properties. If continuous flow had been used, then observers could potentially have grouped the windows based on the local zero- and first-order spatial derivatives of the velocity field. In particular, the target window could be found by fitting the motion in each window with an affine transformation and then grouping windows with similar fits. Because the target window would be the only window with a unique set of affine parameters, it would be excluded from this group. This approach has proved successful for computers (Adiv, 1985; Wang & Adelson, 1994), but seems less feasible for humans because of our insensitivity to small velocity gradients (Nakayama, 1985). In any event, this approach would fail for the stimuli used here because different sets of affine parameters would be calculated for each of the eight windows in these stimuli.

session, the translation stimuli were used to teach the naïve subjects which key on the computer keyboard corresponded to each of the eight windows. The subjects spent the remaining 40 min of the training session practicing on alternating blocks of spiral and deformation stimuli. The training stimuli had a 2 sec duration and a highly discrepant target (a direction deviation of 180 deg). Auditory feedback was provided on incorrect trials in both the training and experimental sessions.

Two 1-hr experimental sessions were conducted on different days during the week following the training session. The experiment was broken down into blocks of 100 trials: 10 practice trials followed by 90 experimental trials. The translation, spiral and deformation conditions were run in separate but interleaved blocks. Within a block of trials, the different levels of direction deviation were presented in random order. In all, 70 trials were run for each combination of ground pattern and direction deviation. The presentation of the first stimulus was initiated by the subject, subsequent stimuli were presented 500 msec after a response. The stimulus duration was 70 frames (700 msec).

Subjects. In total, the author and three paid observers participated in the three experiments reported here. The three paid observers were recruited from the general student population at the University and had no previous experience as research subjects.

Results and discussion

Before considering the results of this experiment, recall the prediction set out in the Introduction. If observers used only information about velocity differences to segregate the target motion in these stimuli, then the spiral and deformation conditions should produce the same level of performance, since the magnitude of the velocity differences across neighboring windows is the same for these two stimuli. Figure 5 shows the results for the spiral and deformation conditions with the percentage of correct responses plotted as a function of the direction deviation of the target window. When the direction deviation was zero, observers had a 1 in 8 chance of guessing the target window. As the magnitude of the direction deviation increased, performance on the spiral and deformation conditions diverged with performance on the spiral condition (filled circles), clearly superior to performance on the deformation condition (open circles). In fact, the two naïve observers were essentially unable to find the target in the deformation stimuli for all levels of direction deviation. This marked difference in performance on the spiral and deformation conditions indicates that observers are using more than just local information to find the target.

As expected, observers found it easiest to locate the target motion when it was added to a translation pattern (filled squares). In the translation stimuli the motion was identical in all the windows except for the target window, and so even the simplest of segregation strategies could be used to find the target in these displays. Since the translation condition differed from the other two conditions in both its local and global motion characteristics, the superior performance in this condition could be due to segregation processes acting at a local or a global level.

The conclusion of the main experiment, that motion segregation involves a global process, is the central message of the paper. The remaining two experiments were less theoretically motivated, and instead test some of the assumptions that were implicit in the design of the main experiment. The first of these experiments examines whether spiral patterns should be considered a homogeneous set or whether there are variations in performance across these patterns. The second experiment examines whether the difference in performance on spiral and deformation patterns is only observed when the patterns are confined to windows.

SUBSIDIARY EXPERIMENTS: SUB-EXPERIMENT 1

In the main experiment, the spiral patterns were treated as a single condition, and the results were pooled across this condition. However, it is quite possible that not all spiral patterns produced the same level of performance. One might expect observers to perform best on the flow patterns that they see most often. If, as is often argued, we see expansion patterns more frequently than contraction patterns because we typically move forwards through the world, then observers might perform better with expansion stimuli than contraction stimuli.

It seemed impractical to use the data from the first experiment to compare performance across spiral patterns, since this experiment involved 360 different spiral patterns. Thus, a second experiment was conducted. In this subsidiary experiment only eight spiral patterns were used: positive and negative expansions, positive and negative rotations, and four mixed spirals composed of expansions and rotations of equal magnitude.

Methods

The methods were the same as those used in the main experiment except that the directions assigned to the first window ranged from 22.5 to 337.5 deg in 45 degree increments. As before, the observer's task was to locate the one window with a direction of motion that deviated from the global motion pattern. The magnitude of this direction deviation was fixed at 45 degrees for TC and 22.5 deg for MB.* Seventy trials were run for each of these patterns, and within a block of trials the eight patterns were presented in random order.

One subject, JD, repeated the experiment with six direction deviations: 22.5, 33.75, 45, 56.25, 67.5, and 90 deg. Each level of deviation was run in a separate block of trials, with blocks for the different levels randomly intermixed.

*This experiment was conducted several months after the main experiment and the following experiment. In the intervening months, subjects TC, MB and JD participated in several related experiments. It is probably because of this additional practice that subjects required smaller direction deviations in this experiment than might be predicted from the results of the main experiment.

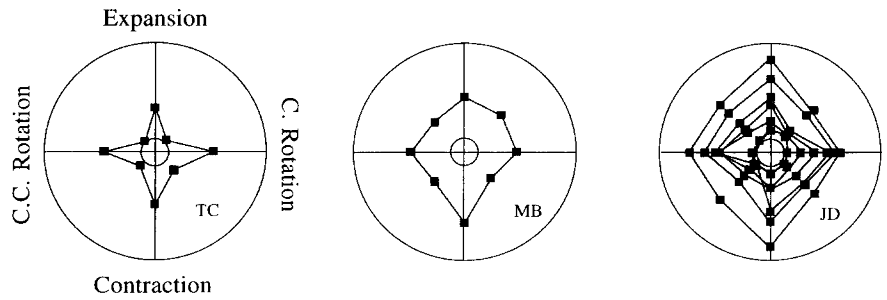

FIGURE 6. Percentage of trials in which the observer correctly located the target window (radius) plotted against the type of spiral pattern (angle). The inner circle in each plot shows 12.5% correct, or chance, and the outer circle shows 100% correct. Two observers performed this experiment with a single level of direction deviation (45 deg for JD and TC, 22.5 deg for MB). One subject, JD, ran the experiment with six levels of direction deviation.

Results and discussion

The polar plots in Fig. 6 show the percentage of correct responses (radius) plotted against the type of spiral (angle). Two graphs show performance for a single direction deviation, and the third graph shows performance on six levels of direction deviation. If observers performed equally well on all the spiral patterns, then the data would fall on a circle and clearly they do not. All three subjects performed better on pure expansions and pure rotations than on the mixed spirals.

Performance in these experiments was measured as the percentage of trials in which the observer selects the

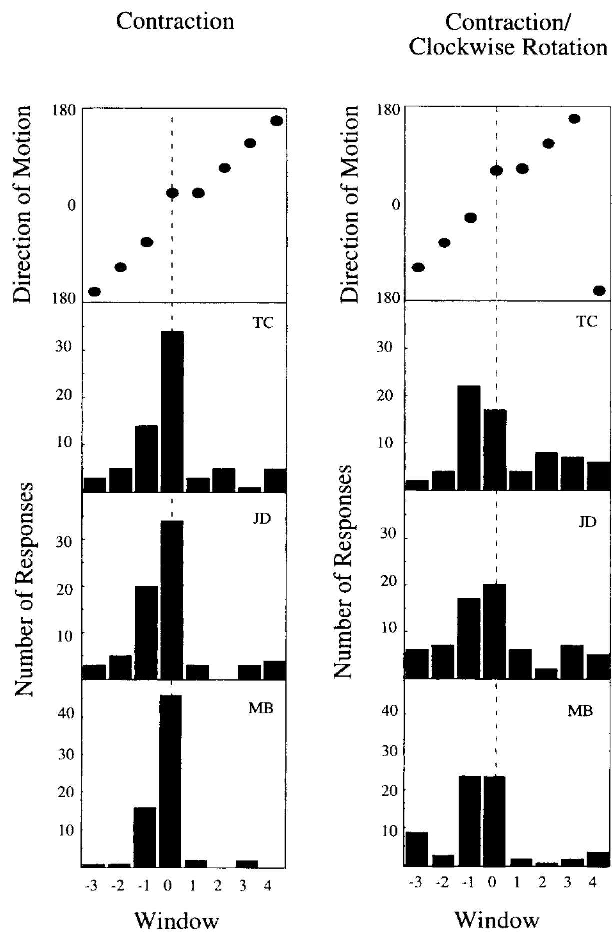

FIGURE 7. Histograms showing the number of times the observers selected each window, where the windows are identified by their distance from the target. The column of graphs on the right shows the data from a mixed spiral (contraction/clockwise rotation), the column on the left shows the data from a pure spiral (contraction). Trials in which the observer selected the target's similar neighbor were collected in bin 1, and trials in which the observer selected the target's dissimilar neighbor were collected in bin -1. Correct trials were collected in bin 0. Note that on a disproportionate number of error trials, observers selected the target's dissimilar neighbor.

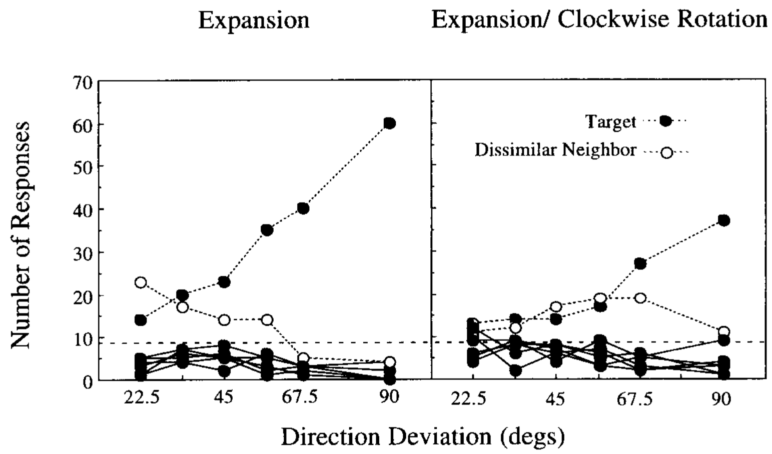

FIGURE 8. Plots of the number of times an observer (JD) selected each window, defined relative to the target window, as a function of the target's direction deviation. The left plot shows the data from a pure expansion, the right from a mixed spiral.

target window. This measure would completely characterize performance if, on error trials, the observers randomly selected the other seven windows. However, as the histograms in Fig. 7 show, the pattern of errors was not random. These histograms were generated by first sorting the data by the type of spiral, and then by the difference between the selected window and the target window, and then finally, by the direction of the target deviation. This final sort was necessary because of the two windows neighboring the target: one window contained motion similar (or identical) to the target, and the other contained dissimilar motion. For example, when the target deviation was +45 deg, the target was identical to its counterclockwise neighbor, and differed by 90 deg from its clockwise neighbor. When the target deviation was −45 deg, the opposite relationship held. Trials in which the observer selected the target's similar neighbor were collected in bin 1, and trials in which the observer selected the target's dissimilar neighbor were collected in bin −1. Correct trials were collected in bin 0. The histograms in Fig. 7 show representative data for an "easy" pattern, a contraction, and a "difficult" pattern, a mixed spiral. The pattern of errors was highly consistent across all three observers and all eight spirals: on a disproportionate number of error trials, the observers selected the target's dissimilar neighbor.

The finding that the distribution of error responses was not random raises the concern that the method used here produces a biased measure of performance. It is conceivable that observers detect the target as readily in mixed spirals as in pure spirals, but they mislocate the target more often in mixed spirals. If so, then a task which required less precise localization (e.g., identifying which side of the display contained the target), might have produced qualitatively different results. To determine whether there were consistent differences in the pattern of responses to different stimuli, I examined the proportion of target responses (bin 0) to dissimilar neighbor responses (bin −1) as a function of the direction deviation of the target. Figure 8 shows the distribution of JD's responses for an expansion stimulus and a mixed spiral stimulus. Although the two graphs differ when superimposed, they fall into register when the expansion data are shifted leftward relative to the mixed spiral data. That is, when the number of target responses is matched for the two stimuli, the number of dissimilar neighbor responses also matches. This relationship held for the other six spiral stimuli as well, suggesting that the task of selecting the target window is an unbiased measure of performance.

One explanation for why observers often select the target's dissimilar neighbor is that they locate the target by locating the largest velocity difference in the stimulus. However, if observers can locate the target in this way, then they should have been able to locate the target in the deformation stimuli of the main experiment. An alternative explanation for the observed pattern of errors is suggested by the subjective reports of the observers. When presented with a large direction deviation, the observers reported that the target appeared to "pop-out" from the background pattern and they could easily locate the target window. When presented with a small direction deviation, the observers reported seeing a very regular pattern and they performed at chance levels. Between these two extremes were displays in which the target did not pop-out from the background pattern, but the pattern appeared distorted. Observers had no trouble identifying which side of the display was distorted, but they could not always identify the window that was responsible for this distortion. When presented with these intermediate displays, an observer might direct her attention to the distorted region of the pattern, notice the large velocity difference between the target and its dissimilar neighbor, and select either window. By this account, a global pattern matching process guides a local process in which neighboring windows are compared.

Regardless of the explanation for the pattern of errors, the subjective reports of the observers indicate that the methods used here are an imperfect measure of segregation. It is possible to perform this task reliably, but probably not perfectly, without segregating the target from the background; when the target is incorporated into

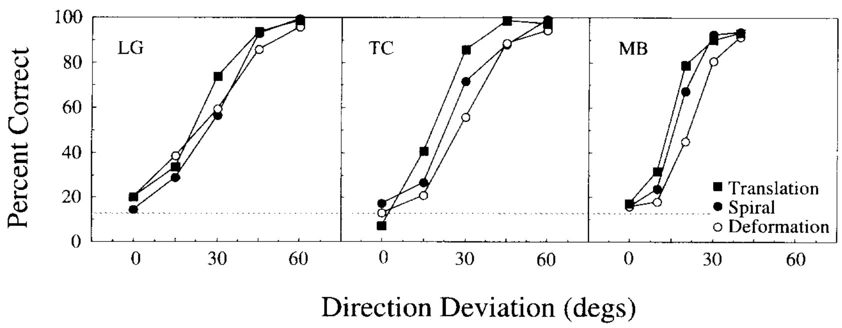

FIGURE 9. Percentage of trials in which the subject correctly identified the target window plotted against the direction deviation of the target. Each panel shows the performance of one subject on the three conditions. These data, which were obtained with dense flow fields, should be compared with the data in Fig. 5, which were obtained with patchy flow fields.

the global pattern, observers may still locate the distortion it produces in the pattern. Similarly, the task used here is not a clear indicator of grouping. An observer may be unable to locate the target either because she is unable to group the background windows into a single pattern, or because she grouped the target into this global pattern. That is, the target may not segregate from the background because grouping is too weak, or because it is, in a sense, too strong. So while this experiment shows that it is easier to locate a discrepancy in a "pure" spiral than in a mixed spiral, it does not reveal the cause of this performance difference.

SUB-EXPERIMENT 2

In designing the first experiment, I had assumed that local velocity discontinuities are a very potent cue for segregation and that it would be necessary to reduce this cue if a global segregation process were to be revealed. So rather than embedding the target window in a dense flow field that changed smoothly over space, the target window was added to a patchy flow field in which the motions were confined to windows. The patches of the flow field were all consistent with the same global motion pattern, but neighboring patches had very different velocities. These large velocity differences were expected to obscure the velocity differences surrounding the target. I examined whether this precaution was necessary by repeating the main experiment without windows. In this second experiment, the target was embedded in a uniformly dense flow field.

Methods

This experiment involved the same three global patterns that were used in the first experiment, that is, spirals, deformations and translations. But whereas before velocity changed in discrete steps across the display (that is, velocity was constant within a window and changed only across windows), in this experiment velocity changed according to a continuous function. The three global patterns were generated as follows.

Spiral. On the first frame of each cinematogram, 580 dots were assigned random spatial locations (x,y) within a 9.8 deg v.a. square. The x and y components of each dot's velocity (V_x and V_y) were then calculated on a frame-by-frame basis using the following equations: $V_x = ax + by$ $V_y = -bx + ay$, where a and b were chosen randomly on each trial with the only constraint being that the sum of their squares equaled 0.61. This constraint ensured that the stimulus speed would be 2.5 deg/sec at an eccentricity of 3.2 deg v.a.

Deformation. The procedure for generating deformation stimuli was identical to that for the spiral stimuli, except that the signs of a and b were reversed in the second equation: $V_x = ax + by$ $V_y = bx - ay$.

Translation. For the translation displays, V_x and V_y were constant across the display ($V_x = c$, $V_y = d$), where c and d were chosen randomly on each trial, with the only constraint being that the sum of their squares equaled 6.25. Thus, in these stimuli direction varied randomly across trials, but speed was fixed at 2.5 deg/sec.

Embedded in these continuous global patterns was a square window, 1.22 deg.v.a. wide, filled with a target motion. The target window appeared in one of eight locations arranged at 45-deg intervals around an imaginary circle with a radius of 3.2 deg v.a. Again the subject's task was to identify the location of the target window, and although the eight possible locations were not marked in the display, subjects had no difficulty in making this identification. The frame-to-frame displacements of the target dots were calculated using the equations given above, but the direction of the displacement was rotated by an angular amount equal to the direction deviation for that trial. On average, nine dots fell within the target window. Dots were not permitted to cross the window's boundaries: when a background dot entered the target window it was replotted in a random location, and, similarly, when a target dot entered the background it was replotted elsewhere in the display.

The dots in these displays had limited lifetimes: after moving with a consistent trajectory for 8 frames, each dot was replotted at a random location. The dot-lifetimes were staggered so that on each frame, approximately 1/8 of the dots were assigned random positions. These limited lifetime displays seemed to twinkle as dots appeared and disappeared at random locations throughout the display.

Limited lifetime dots were used for two reasons. First, casual inspection of the displays indicated that the random dot disappearances throughout the display made the dot disappearances around the target window inconspicuous. This observation was confirmed with stimuli having a 0 direction deviation. In these stimuli, dots were prevented from crossing the boundary of the target window, but the target's motion was consistent with the background pattern. Observers performed at chance when the target was defined only by the disappearances along its boundary (Fig. 9). The second reason for using limited-lifetime dots was that by assigning each dot a new random location every 80 msec, dot density remained roughly uniform across the display. Otherwise, when the initially random pattern of dots was subjected to an expanding flow field, the density of the dots might have become noticeably non-uniform. This non-uniformity in dot density could affect the search for the target, especially if the density of the target window began to deviate from its immediate surround. Thus, the use of limited lifetime dots made it unlikely that observers were using a dot disappearance or a dot density cue to locate the target window.

To summarize, the stimuli used in this last experiment and the main experiment differed in the following ways. Rather than displaying 240 dots in eight windows, 580 dots were scattered uniformly over a square area, 9.8 deg v.a. wide. In this experiment, the velocities of the dots were calculated on a dot-by-dot basis to produce a continuously varying global flow pattern. This differs from the main experiment, in which all of the dots within a window were assigned the same velocity and so only approximated a continuously varying flow pattern. The dots in this experiment had a limited lifetime of 8 frames. The width of the square target window in this experiment was increased from 0.8 to 1.22 deg so that an average of nine dots would fall within the window. The target window was less eccentric (3.2 deg v.a. rather than 4 deg v.a), however, the speed at the center of the target window was still 2.5 deg/sec.

Results and discussion

This last experiment was a variation on the main experiment in which the primary difference was that motions were distributed throughout the display rather than within windows. To generate these continuous displays, however, other aspects of the stimulus had to be modified, the most notable modification being the use of limited lifetime dots. These secondary modifications would be expected to produce a modest difference in performance between the main experiment and this experiment. The observed differences, however, were quite striking (Fig. 9).

Note first that the range of direction deviations used in the main experiment (Fig. 5) was three times that used here. Confining the motion to windows makes the search task considerably more difficult. A second striking difference between the two figures is that the variation across conditions that was so conspicuous in the main experiment is greatly reduced, if not eliminated, in this experiment. The windows appear to be the key to revealing the large effect that the global motion pattern can have on the search for a discrepant motion.

The different pattern of results in the main experiment and this final experiment suggests that observers may use a different segregation process when the target appears in a continuous flow field than when it appears in a patchy flow field. There were large, local velocity discontinuities surrounding the target in the continuous displays of this experiment, but not in the main experiment. These discontinuities would provide strong input to a motion boundary detector which could then signal the location of the target. Since the local velocity discontinuity is matched in the spiral and deformation conditions, a motion boundary detector should work equally well on these two patterns.

GENERAL DISCUSSION

The central finding of this study is that when the dots of the global flow stimuli were confined to windows, subjects were better able to locate a discrepant motion in a spiral pattern than in a deformation pattern. This discrepancy in performance is of interest because the spiral and deformation patterns were matched in their component velocities and velocity differences. With stimuli so matched, the predictions of a segregation process based on boundary detection or region-growing are straightforward. These strategies are discussed below.

After making local measurements of various image properties across a scene, the visual system presumably segregates these measurements into meaningful regions.* One segregation strategy is to detect spatial discontinuities in the measurements of some property and to form region boundaries along these discontinuities. As mentioned earlier, there is evidence that the human visual system applies this strategy to local motion measurements. However, if observers in this experiment had located the target by locating the largest velocity difference in the stimulus, then they should have performed equally well on the spiral and deformation stimuli and they did not. A second, complementary, segregation strategy involves grouping locations with similar measurements. This strategy, termed "region growing," is commonly applied in computer vision and generally involves collecting measurements from across the image in a histogram and then dividing the modes of the resulting frequency distribution. If spatial proximity is not factored into the grouping process, then this

*Admittedly, the notion of meaningful regions is ill-defined and depends in part on the task at hand. For example, if the goal is to find objects that are moving in the environment, then these regions correspond to surfaces that are moving rigidly with respect to one another, regardless of surface attitude or depth. In contrast, if the goal is to recover surface layout then these regions would correspond to continuous surfaces, and thus surfaces at different depths would be segregated even if they had the same 3D motion.

histogram would be two-dimensional, to account for the two dimensions of velocity. If preference is given to grouping spatially proximal motions, then the histogram would be four-dimensional to accommodate spatial position. It is simple to apply this clustering approach to the stimuli used here if one considers a four-dimensional histogram in which velocity is parametrized by direction and speed and spatial position is parametrized by angle and eccentricity. Since window eccentricity and dot speed were constant within and across stimuli, the only interesting slice through these histograms is the slice showing direction of motion vs window angle, that is, Fig. 4. It is obvious from these plots that any standard clustering technique would treat the spiral and deformation stimuli equivalently.

If the segregation strategies of boundary detection and region growing cannot explain the observed difference between spiral and deformation patterns, is there a segregation strategy that can? The two segregation strategies discussed above are based only on the assumption that the motions arising from one surface vary smoothly across space. While this assumption is generally true, it does not exploit the regularities that exist in many naturally occurring flow fields. For example, as we translate forward along the line of sight, the resulting optic flow field is an expanding pattern. The visual system may have a specialized segregation strategy than can take advantage of this regularity: namely, if several motions in the image are consistent with an expanding pattern then they may be grouped and so segregated from motions that are inconsistent with this pattern. Thus, in addition to the general purpose segregation strategies mentioned above, the visual system may test a set of motion models against the stimulus. If one of these models can explain a significant amount of the image, then this model is used to organize the scene. Of course, this model-based strategy can only explain the present set of results if the visual system has a model of spirals but not one for deformations. This account of the results is clearly circular, but it does have independent support. Numerous neurophysiological studies of area MST in the monkey brain have demonstrated the existence of cells that are specifically tuned to global spiral patterns, while little evidence has been found for cells tuned to global deformation patterns (see, for example, Saito *et al.*, 1986; Duffy & Wurtz, 1991a,b; Lagae *et al.*, 1994; Graziano *et al.*, 1994).

Before accepting the idea that the visual system has a model for spirals but not for deformations, we should consider the possibility that the visual system actually does have a model of deformations, but it does not correspond to the stimuli used here. The flow fields used in these experiments were time-invariant: the pattern of motions at the end of the stimulus presentation was the same as in the beginning. Natural flow fields typically evolve over time. Consider, for example, a real-world situation that could produce a flow field similar to pure deformation: an observer picking up a book and turning it over to examine its spine. As the observer brings the book closer, the image of the book expands, but as the observer simultaneously rotates the book, its image contracts in a direction orthogonal to the axis of rotation. If these two motions occurred at the correct rate, then at one instant the flow field would be similar to a pure deformation. It is critical to note, however, that the flow field generated by the book resembles a pure deformation at only one instant in time. As the book rotates, its orientation relative to the observer changes and this produces a significant change in the flow field. Thus, the observers in this experiment saw something that observers of the real world do not see: a pure deformation that is time-invariant. The temporal characteristics of such a stimulus are at odds with its spatial characteristics, and this surely explains why the deformation stimuli looked non-rigid.

While the use of such time-invariant flow fields may be inappropriate for deformation stimuli, they are quite appropriate for patterns composed of expansions and rotations. Rotation patterns can be produced by rotating a frontoparallel plane about the line of sight. This motion does not cause the plane to change its orientation relative to the observer and so the resulting flow field is time-invariant as long as the motion of the plane is constant. Expansion patterns are produced when a plane translates along the line of sight. This motion too does not produce a change in the plane's orientation over time and so the flow pattern does not change. As the plane approaches, however, the rate of expansion increases and this produces a change in speed. There is abundant evidence, however, that humans are quite insensitive to gradual changes in speed (Gottsdanker, 1965; Snowden & Braddick, 1991), and so one might expect that for stimuli of limited duration, the visual system may not discriminate between an "ecological" expansion, one that accelerates over time, and the unecological expansions used here.

To summarize, the present results provide strong evidence for the existence of a global process in motion segregation. One process that could potentially explain the results involves fitting the stimulus with a global motion pattern and segregating motions inconsistent with this pattern. Since observers were able to locate an inconsistent motion much more reliably when it was presented in a spiral pattern than when it was presented in a deformation pattern, the study would seem to imply that the set of global motion patterns includes spirals but not deformations. However, this conclusion should be accepted with caution. As noted earlier, failure to locate the target in these experiments could be due to a failure to group the background windows into a global pattern or to the assimilation of the target into the global pattern. To decide between these possibilities, it will be necessary to devise an independent measure of global pattern detection for these stimuli. Further, the deformation patterns used in these experiments were highly unecological in that they did not change over time. If deformation patterns are among the motion models that the visual system recognizes, these patterns would most likely have a temporal aspect as well as a spatial aspect. Thus, the present set of experiments provides evidence for a global segregation process, but further experiments will be necessary to characterize this process.

REFERENCES

Adiv, G. (1985). Determining three-dimensional motion and structure from optical flow generated by several moving objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-7*, pp. 384–401.

Banton, T. & Levi, D. M. (1993). Spatial localization of motion-defined and luminance-defined contours. *Vision Research, 33*, 2225–2237.

Bergen, J. R., Burt, P. J., Hingorani, R. & Peleg, S. (1992). A three-frame algorithm for estimating two-component image motion. *IEEE Transactions of Pattern Analysis and Machine Intelligence, 14*, 886–896.

Black, M. J. & Anandan, P. (1993). A framework for the robust estimation of optic flow. *Proceedings of the International Conference on Computer Vision, ICCV-93, 93*, 231–236.

Cavanagh, P. (1989). Multiple analysis of orientation in the visual system. In D. M. Lam & C. D. Gilbert (Eds), *Neural mechanisms of visual perception* (pp. 261–279), Cambridge, MA: MIT Press.

Darrell, T. & Pentland, A. (1991). Robust estimation of a multi-layered motion representation. In Proc. IEEE Workshop on Visual Motion, Princeton, NJ, pp. 173–178.

Duffy, C. & Wurtz, R. H. (1991a) Sensitivity of MST neurons to optic flow stimuli, I. A continuum of response selectivity to large field stimuli. *Journal of Neurophysiology, 65*, 1329–1345.

Duffy, C. & Wurtz, R. H. (1991b) Sensitivity of MST neurons to optic flow stimuli, II. Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology, 65*, 1346–1359.

Freeman, T. C. & Harris, M. G. (1992). Human sensitivity to expanding and rotating motion: effects of complementary masking and directional structure. *Vision Research, 32*, 81–87.

Gottsdanker, R. M. (1965). The ability of human operators to detect acceleration of target motion. *Psychological Bulletin, 53*, 477–487.

Graziano, M. S. A., Andersen, R. A. & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. *Journal of Neuroscience, 14*, 54–67.

Halpern, D. L. (1991). Stereopsis from motion-defined contours. *Vision Research, 31*, 1611–1617.

Irani, M., Rousso, B. & Peleg, S. (1992). Detecting and tracking multiple moving objects using temporal integration. In G. Sandini (Ed)., Proceedings of the Second European Conf. Computer Vision, ECCV-92 (Vol. 588 of LNCS series, pp. 282–287). Berlin: Springer-Verlag.

Lagae, L., Maes, H., Raiguel, S., Xiao, D. K. & Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *Journal of Neurophysiology, 71*, 1597–1626.

Lappin, J. S., Norman, J. F. & Mowafy, L. (1991). The detectability of geometric structure in rapidly changing optical patterns. *Perception, 20*, 513–528.

Marr, D. (1982). *Vision* (p. 51). San Francisco: W. H. Freeman & Co.

Morrone, M. C., Burr, D. C. & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature, 376*, 507–509.

Nakayama, K. (1985). Higher order derivatives of the optical velocity vector field: limitations imposed by biological hardware In D. Ingle, M. Jeannerod & D. Lee (Eds), *Brain mechanisms and spatial vision*. Holland: Martinus Nijoff.

Narwot, M., Shannon, E. & Rizzo, M. (1996). The relative efficacy of cues for two-dimensional shape perception. *Vision Research, 36*, 1141–1152.

Petersik, J. T., Hicks, K. I. & Pantle, A. J. (1978). Apparent movement of successively generated subjective figures. *Perception, 10*, 563–572.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience, 6*, 145–157.

Snowden, R. & Braddick, O. (1991). The temporal integration and resolution of velocity signals. *Vision Research, 31*, 907–914.

Wang, J. Y. A. & Adelson, E. H. (1994). Representing moving images with layers. *IEEE Transactions on Image Processing Special Issue: Image Sequence Compression, 3*, 625–638.

Warren, W. H. & Hannon, D. J. (1988). Direction of self-motion is perceived from optic flow. *Nature, 336*, 162–162.

Warren, W. H., Morris, M. W. & Kalish, M. (1988). Perception of translational heading from optic flow. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 646–660.

Yuille, A. & Grzywacz, N. M. (1997). A theoretical framework for visual motion. In T. Watanabe (Ed.), *High-level motion processing—an interdisciplinary approach*. North Holland Press, in press.

Acknowledgements—This work was supported by NEI RO1 EY11032. I would like to acknowledge David Jauch and Hany Farid for their invaluable assistance. I am also grateful to David Knill, Jacob Nachmias, Paul Schrater, Eero Simoncelli, and two anonymous reviewers for their advice and suggestions.